INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Zero distributions via orthogonality

L. Baratchart — R. Küstner — V. Totik

N° 5291

Août 2004

Thème NUM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--5291--FR+ENG

Zero distributions via orthogonality

L. Baratchart , R. Küstner , V. Totik*

Thème NUM — Systèmes numériques
Projet Apics

Rapport de recherche n° 5291 — Août 2004 — 34 pages

Abstract: We develop a new method to prove asymptotic zero distribution for different kinds of orthogonal polynomials. The method directly uses the orthogonality relations. We illustrate the procedure in four cases: classical orthogonality, non-Hermitian orthogonality, rational approximation of Markov-type functions and its non-Hermitian variant. In the last three cases, the results are first of this kind.

Key-words: Function theory, Potential Theory, Orthogonal Polynomials, Non-Hermitian Orthogonal Polynomials, Rational Approximation, Meromorphic Approximation.

Dedicated to Zbigniew Ciesielski on his 70th birthday

Submitted to *Annales de l'Institut Fourier*

* Supported by NSF grant DMS-0097484 and by OTKA T/034323, TS44782

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 —Télécopie : +33 4 92 38 77 65

Distribution de zéros et relations d'orthogonalité

Résumé : Ce rapport présente une nouvelle méthode pour déterminer la distribution asymptotique des zéros de polynômes orthogonaux sur un segment, qui procède directement à partir des relations d'orthogonalité. Nous l'illustrons dans quatre cas de figure : orthogonalité classique, non-Hermitienne, orthogonalité en approximation rationnelle pour les fonctions de Markov, ainsi que pour les intégrales de Cauchy de mesures complexes. Les résultats dans les trois derniers cas sont les premiers de ce type.

Mots-clés : Théorie des fonctions, Théorie du potentiel, Polynômes orthogonaux, Polynômes orthogonaux non-Hermitiens, Approximation rationnelle, Approximation méromorphe.

1 Introduction

Let μ be a positive measure with infinite compact support $S \subset \mathbf{R}$, and consider the corresponding monic orthogonal polynomials $q_n(t) = t^n + \cdots$,

$$\int q_n(t)\, t^k \, d\mu(t) = 0, \qquad k = 0, \ldots, n-1. \tag{1}$$

Under quite weak hypotheses on the measure μ, it is true [19, Theorem 2.2.1] that the zero distribution of the polynomials q_n is asymptotically equal to the equilibrium distribution ω_S of S for the logarithmic potential. The proofs in the literature usually go through estimates on the norm of the polynomials. In what follows, we present a new approach that uses directly the orthogonality relations (1). Under mild assumptions this approach generalizes to non-Hermitian orthogonality, that is to the case where μ gets replaced by some complex measure λ; this cannot be said of classical proofs. See further motivation for and connection with non-Hermitian orthogonality in Section 6.

Another feature of the present approach, is that it applies to other kinds of orthogonality relations, like those arising from best rational or meromorphic approximation with free poles. For example, if the support of λ lies in the unit disk, the denominator of a best rational approximant of degree n to the Cauchy transform of λ in L^2 of the unit circle satisfies [7, 3]:

$$\int \frac{q_n}{\widetilde{q}_n^2}(t)\, t^k \, d\lambda(t) = 0, \qquad k = 0, \ldots, n-1, \tag{2}$$

where $\widetilde{q}_n$ is the reciprocal polynomial associated with q_n (see Section 4). Although (2) looks like ordinary orthogonality with varying weight $\widetilde{q}_n^{-2}$, it is not so because this weight itself depends on q_n. When λ is a positive measure on $[a, b] \subset (-1, 1)$ satisfying the Szegő condition, then the asymptotic zero distribution of the polynomials q_n in (2) is known (and much beyond, see [2, 5]), but still our result deals with less regular situations. And when λ is complex, the theorem we prove is first of this kind. The method also allows one to handle the case of a strongly convergent varying weight which is relevant in meromorphic approximation [3].

In this work we shall extensively use potential theory. For concepts like equilibrium measure, logarithmic potential, capacity, balayage, and the basic theorems of potential theory, the reader may consult some recent texts such as [14], [15] or the appendix of [19].

2 Real orthogonal polynomials

In this section, we let μ to be a positive Borel measure with support $S \subset \mathbf{R}$. We shall assume that μ is sufficiently thick, namely that there are $c, L > 0$ such that, for all $0 < \delta < 1$ and $x \in S$, we have

$$\mu([x-\delta, x+\delta]) \geq c\delta^L. \tag{1}$$

Let ν_n be the normalized counting measure on the zeros of the polynomials q_n satisfying (1), namely the discrete probability measure having equal mass at each of the zeros (these are simple). The theorem below asserting the asymptotic behaviour of ν_n is not new see [19, Theorems 2.2.1 and 4.2.5]), but our method of proof will serve as a model for more general situations to come.

Theorem 2.1 *Suppose that the support S of μ is regular with respect to the Dirichlet problem in $\overline{\mathbf{C}} \setminus S$, and that (1) holds. Then ν_n tends to the equilibrium measure ω_S of S in the weak* topology as n tends to ∞.*

Proof. Let ν be a weak* limit point of $\{\nu_n\}$, say $\nu_n \to \nu$ as $n \to \infty$, $n \in \mathbf{N}_1$. It is well-known (and elementary to check see [19, Lemma 1.1.3]), that the zeros of the orthogonal polynomials lie in the convex hull of S, that they are simple, and that each subinterval of $\mathbf{R} \setminus S$ can contain at most one of them. Hence ν is supported on S and it has total mass 1. *We claim:* it is enough to prove that there is a constant D such that the logarithmic potential U^ν of ν equals D quasi-everywhere on S. Indeed, the lower semi-continuity of U^ν implies then that $U^\nu(x) \leq D$ for all $x \in S$. Thus, ν has finite logarithmic energy, and, in particular, every set of zero capacity has zero ν measure. Now, integrating the equality "$U^\nu(x) = D$ for quasi-every $x \in S$" agains ω_S and interchanging the order of integration, we get that

$$D = \int U^\nu d\omega_S = \int U^{\omega_S} d\nu = \log \frac{1}{\operatorname{cap}(S)},$$

and then integrating the same equality against ν yields that ν has the same logarithmic energy as ω_S. By uniqueness of the equilibrium measure (c.f. [15, Theorem I.3.1]) it follows that $\nu = \omega_S$ and, since this is true of every weak* limit point, the whole sequence $\{\nu_n\}$ converges to ω_S *as claimed.*

Thus, it has left to prove that there is a constant D such that the potential of ν equals D quasi-everywhere on S. Suppose to the contrary that there are d, $\tau > 0$ and two Borel sets $E_- \subset S$ and $E_+ \subset S$ of positive capacity such that

$$U^\nu(x) \leq d - 2\tau \quad \text{for } x \in E_- \qquad \text{and} \qquad U^\nu(x) \geq d + \tau \quad \text{for } x \in E_+$$

(this is the only alternative for U^ν is finite quasi-everywhere on S since it is a superharmonic function on $\mathbf{C}$ which is clearly not identically $+\infty$).

We contend that there is a point $y_0 \in \operatorname{supp}(\nu)$ such that $U^\nu(y_0) > d$. For if not $U^\nu(x) \leq d$ for all $x \in \operatorname{supp}(\nu)$, so by the maximum principle for potentials (see e.g. [15, Corollary II.3.3]) the same inequality holds for all $z \in \mathbf{C}$, a contradiction since on E_+ we have bigger values for U^ν. *This proves our contention.*

According to the principle of descent (see e.g. [15, Theorem I.6.8]), which is valid since the support of ν_n remains in a fixed compact set (namely the convex hull of S), we subsequently have

$$\liminf_{n\to\infty,\ n\in\mathbf{N}_1} U^{\nu_n}(y_n) \geq U^\nu(y_0) > d$$

for any sequence $y_n \to y_0$. Therefore there are $\rho > 0$ and N_1 such that, for $y \in [y_0 - \rho, y_0 + \rho]$ and $n \geq N_1$, $n \in \mathbf{N}_1$, the inequality $U^{\nu_n}(y) \geq d$ holds, which amounts to

$$|q_n(y)| \leq e^{-nd}, \qquad y \in [y_0 - \rho, y_0 + \rho] \tag{2}$$

by the definition of ν_n and of the logarithmic potential. We shall use that this same inequality is true (for sufficiently large n) if $\{q_n\}$ gets replaced by any sequence of monic polynomials $\{q_n^*\}$ having the same asymptotic zero-distribution ν and whose zeros are uniformly bounded, for these were the only facts we used in deriving (2). This is true even if q_n^* does not have exact degree n (the counting measure of the zeros being still normalized with n), but merely degree $(1+o(1))n$.

In another connection, the lower envelope theorem (see [15, Theorem I.6.9]) implies that for quasi-every $x \in E_-$ we have

$$\liminf_{n\to\infty,\ n\in\mathbf{N}_1} U^{\nu_n}(x) = U^{\nu}(x) \leq d - 2\tau,$$

and since the logarithm of $|q_n(x)|^{-1/n}$ stands on the left, we deduce that for some subsequence $\mathbf{N}_2 \subset \mathbf{N}_1$ and for sufficiently large $n \in \mathbf{N}_2$, say for $n \geq N_2$

$$M_n := \max_{x\in S} |q_n(x)| \geq e^{-n(d-\tau)}. \tag{3}$$

Our assumption that S is regular means: the Green function $g_{\mathbf{C}\setminus S}(z)$ of $\mathbf{C}\setminus S$ with pole at ∞ is such that

$$\lambda(\varepsilon) := \max_{\mathrm{dist}(z,S)\leq 2\varepsilon} g_{\mathbf{C}\setminus S}(z)$$

tends to zero as $\varepsilon \to 0$. Now, according to the Bernstein-Walsh lemma [20, p. 77], we have that

$$|q_n(z)| \leq M_n \exp(n g_{\mathbf{C}\setminus S}(z)) \leq M_n e^{n\lambda(\varepsilon)}$$

if $\mathrm{dist}(z,S) \leq 2\varepsilon$, hence for $\mathrm{dist}(z,S) \leq \varepsilon$ we obtain from Cauchy's formula the estimate:

$$|q_n'(z)| \leq \frac{M_n e^{n\lambda(\varepsilon)}}{\varepsilon}$$

where $q_n'(z)$ indicates the derivative. Thus, if we let $x_n \in S$ be a point where $|q_n|$ attains its maximum on S and if $x \in S$ satisfies:

$$|x - x_n| \leq \frac{\varepsilon}{2e^{n\lambda(\varepsilon)}}, \tag{4}$$

we obtain from the mean-value theorem the estimate:

$$|q_n(x)| \geq M_n - |q_n(x_n) - q_n(x)| \geq M_n - M_n/2 = M_n/2.$$

For fixed $\varepsilon > 0$, noticing since $\lambda(\varepsilon) > 0$ that the interval defined by (4) is contained in $[x_n - \rho/2, x_n + \rho/2]$ when n is large enough, we conclude on invoking (1) and (3) that

$$\begin{aligned} \int_{[x_n-\rho/2,x_n+\rho/2]} |q_n|^2 d\mu \geq \left(\frac{M_n}{2}\right)^2 c\left(\frac{\varepsilon}{2e^{n\lambda(\varepsilon)}}\right)^L &\geq \frac{c}{4}\left(\frac{\varepsilon}{2e^{n\lambda(\varepsilon)}}\right)^L e^{-2n(d-\tau)} \\ &\geq e^{-2nd+n\tau} \end{aligned} \tag{5}$$

for $n \in \mathbf{N}_2$ and, say, $n \geq N_3 = N_3(\varepsilon)$, provided $\varepsilon > 0$ was chosen so small that $\lambda(\varepsilon)L < \tau$.

Observe now that, for sufficiently large $n \in \mathbf{N}_1$, there are arbitrarily many zeros of q_n lying in $[y_0 - \rho/2, y_0 + \rho/2]$ because y_0 lies in the support of the limit measure ν. In particular there will be two such zeros, say $\alpha_{n,1} < \alpha_{n,2}$, as soon as $n \in \mathbf{N}_1$ is sufficiently large, say $n \geq N_4$. For $n \in \mathbf{N}_2$ with $n \geq \max(N_1, N_2, N_3, N_4)$, consider the polynomial $P_{n-2}(t) = q_n(t)/(t-\alpha_{n,1})(t-\alpha_{n,2})$ of degree $n-2$. Obviously, $q_n(t)P_{n-2}(t)$ is nonnegative outside $(\alpha_{n,1}, \alpha_{n,2})$. Moreover, since x_n was a maximum place for $|q_n|$ on S, it follows from (2) and (3)) that $x_n \notin [y_0 - \rho, y_0 + \rho]$ so that $[x_n - \rho/2, x_n + \rho/2] \cap (\alpha_{n,1}, \alpha_{n,2}) = \emptyset$. Consequently, by (5), we obtain on the one hand for large $n \in \mathbf{N}_2$:

$$\int_{t \notin (\alpha_{n,1}, \alpha_{n,2})} q_n(t) P_{n-2}(t) d\mu(t) \geq \frac{1}{(\mathrm{diam}(S))^2} e^{-2nd+n\tau}. \tag{6}$$

On the other hand, the limit distribution of the zeros of the polynomials P_{n-2}, $n \in \mathbf{N}_2$ is again ν and these zeros lie in the convex hull of S so that, as pointed out after (2), we also have:

$$|P_{n-2}(y)| \leq e^{-nd}, \qquad y \in [y_0 - \rho, y_0 + \rho],$$

for sufficiently large $n \in \mathbf{N}_2$. We thus obtain from this and (2) the estimate:

$$\left| \int_{t \in (\alpha_{n,1}, \alpha_{n,2})} q_n(t) P_{n-2}(t) d\mu(t) \right| \leq C e^{-2nd} \tag{7}$$

with $C = \mu((\alpha_{n,1}, \alpha_{n,2}))$.

But the sum of the two integrals in the left-hand sides of (6) and (7) should be zero by orthogonality, which is clearly impossible because, as an inspection of the right hand sides shows, the first integral is much bigger than the second one for large $n \in \mathbf{N}_2$.

This contradiction proves the theorem.

■

3 Non-Hermitian orthogonal polynomials

Let λ be a complex-valued Borel measure with compact support $S \subset \mathbf{R}$, and consider an associated sequence of monic non-Hermitian orthogonal polynomials namely a sequence $q_n(x) = x^{d_n} + \cdots$, with $d_n \leq n$, such that:

$$\int q_n(t)\, t^k \, d\lambda(t) = 0, \qquad k = 0, \ldots, n-1. \tag{1}$$

Although the orthogonalization process can no longer be used here, the existence of such a sequence $\{q_n\}$ is guaranteed since n linear homogeneous equations in $n+1$ unknowns always have a nontrivial solution. Therefore, there are $a_{n,n}, a_{n,n-1}, \ldots, a_{n,0}$, not all zero, such that

$$\int \left(\sum_{j=0}^{n} a_{n,j}\, t^j \right) t^k d\lambda(t) = 0, \qquad k = 0, 1, \ldots, n-1,$$

and now if a_{n,d_n} is the highest non-vanishing coefficient, then normalization gives (1). This time, however, q_n need not be unique nor have exact degree n; in fact (1) has exactly one monic solution of minimal degree and if there is another solution, say of degree m, then each monic polynomial of degree at most m which is a multiple of the minimal degree solution is in turn a solution [17, lemma 1]. Therefore, there is a certain inaccuracy in the words "a sequence of monic non-Hermitian orthogonal polynomials associated with λ". Hereafter, we simply assume that some monic q_n of degree $d_n \leq n$ satisfying (1) has been chosen for each $n \geq 1$, for the results will not depend on the precise choice of q_n meeting these requirements. Note if the support S of λ is infinite then d_n necessarily goes to infinity with n, as follows easily from the uniform density of polynomials in the space of continuous functions on S. The hypotheses on λ to come will in fact ensure much more, namely they imply that $n - d_n$ remains bounded (see Lemma 3.2). For that very reason, it does not matter in convergence issues whether the zero counting measure ν_n of q_n is formed by putting mass k/d_n or mass k/n to each zero of q_n of multiplicity k. For definiteness, we make the former definition so that ν_n is still a probability measure.

About the measure λ we assume that it is of the form

$$d\lambda(t) = e^{i\varphi(t)} d\mu(t), \tag{2}$$

where φ is a function of bounded variation on the support S of μ, and μ is a positive measure satisfying the properties set forth in Section 2, i.e. its support S is regular with respect to the Dirichlet problem in $\overline{\mathbf{C}} \setminus S$, and (1) holds. An equivalent way to say this is to let μ to be the total variation measure of the complex measure λ and let

$$d\lambda(t) = e^{i\varphi(t)} d\mu(t)$$

be the Radon-Nikodym decomposition of λ with respect to μ (polar form of λ). Thus, μ is a positive measure whose support is the same as the support of λ, and $d\lambda/d\mu(t) = e^{i\varphi}$ is of bounded variation (more precisely it coincides μ-almost everywhere with a function of bounded variation). In fact, it is easy to see that φ and $e^{i\varphi}$ are simultaneously of bounded variation or not on S.

Note that, in the present setting, q_n may have nonreal zeros, so in general, ν_n is not supported on $\mathbf{R}$. However, the next theorem entails that it tends to be so asymptotically, at least in proportion:

Theorem 3.1 *With the preceding assumptions on λ, the measure ν_n tends to the equilibrium measure ω_S of S in the weak* topology as n tends to ∞.*

For the proof of Theorem 3.1, we shall rely on a lemma which is of independent interest. This lemma does not require μ to satisfy (1), nor that S be regular but only that it is infinite. To state the lemma, we need another piece of notation as follows.

Suppose that $S \subseteq \cup_{j=1}^{m} [a_j, b_j]$ where $a_1 < b_1 < a_2 < \cdots < a_m < b_m$. For $\xi \in \mathbf{C}$, let

$$\mathrm{Angle}(\xi, [a, b]) = |\arg(a - \xi) - \arg(b - \xi)| \in [0, \pi]$$

be the angle in which the interval $[a, b]$ is seen from ξ. Subsequently we define

$$\theta(\xi) = \left(\sum_{j=1}^{m} \mathrm{Angle}(\xi, [a_j, b_j]) \right),$$

which is the the total angle in which $\cup_j [a_j, b_j])$ is seen from ξ, and we observe that $\theta(\xi) \in [0, \pi]$ since $\theta(\xi) \leq \mathrm{Angle}(\xi, [a_1, b_m])$. Note also that $\theta(\xi) = 0$ if, and only if $\xi \in \mathbf{R} \setminus \cup_j [a_j, b_j])$, and that $\theta(\xi) = \pi$ if, and only if $\xi \in \cup_j [a_j, b_j])$.

We indicate by

$$V(\varphi) = \sup_{\substack{x_0 < x_1 < \ldots < x_N \\ N \in \mathbf{N},\ x_k \in S}} \sum_{k=1}^{N} |\varphi(x_i) - \varphi(x_{i-1})|$$

the total variation of φ on S.

Lemma 3.2 *Let $q_n(z) = \prod_{k=1}^{d_n} (z - \xi_k)$ be some n-th orthogonal polynomial in the sense of* (1). *Under the sole hypothesis that the support of λ is an infinite compact subset of the real line, it holds with the previous notations that*

$$\sum_{k=1}^{d_n} (\pi - \theta(\xi_k)) + (n - d_n)\pi \leq V(\varphi) + (m-1)\pi, \tag{3}$$

where φ is any argument function for $d\lambda / d|\lambda|$ (i.e. such that $e^{i\varphi(t)} = d\lambda/d|\lambda|(t)$).

It follows from the lemma that the defect $n - d_n$ in the degree of q_n is at most $V(\varphi)/\pi$ and therefore remains uniformly bounded whenever φ can be chosen of bounded variation. To see it, apply the case $m = 1$ of Lemma 3.2, i.e. consider just one interval $[a_1, b_1]$ containing S, and use that $\theta(\xi_k) \leq \pi$ for each k. In particular, when $V(\varphi) = 0$, say λ is positive, we conclude that there is no defect and that all the poles lie in the convex hull of S, which is a well-known result.

Corollary 3.3 *Under the assumptions of Theorem* 3.1, *to every neighborhood U of S there is a constant K_U such that q_n has at most K_U zeros outside U.*

The corollary is an immediate consequence of Lemma 3.2: Select finitely many intervals $[a_j, b_j]$, $j = 1, \ldots, m$ such that

$$S \subseteq \sum_{j=1}^{m} [a_j, b_j] \subset U,$$

and apply (3) noting that a zero outside U contributes to the left hand side by more than a fixed constant that depends only on $\cup_j [a_j, b_j]$ and U.

Proof of Lemma 3.2. We may assume that φ has bounded variation and that it is defined on the whole real line upon extending it linearly to the finite contiguous intervals of S while making it constant and equal to $\varphi(a)$ (resp. $\varphi(b)$) on $(a,+\infty)$ (resp. $(-\infty,b)$) if $[a,b]$ is the convex hull of S; this extension has the same variation as φ. Also, we can suppose that every jump of φ has amplitude at most π, because there are only finitely many discontinuity points where it is not so (by the finiteness of the variation) and then we can always add to $\varphi(t)$ a piecewise constant function with jumps in multiples of 2π to change all jumps larger than π to jumps at most π (in absolute value). This modification can only decrease the variation, and therefore will not affect the validity of (3).

For the proof, we may further assume that q_n has no zero on $\cup_j[a_j,b_j]$. Indeed, putting $d\mu_1(t)=(t-\xi_k)^2d\mu(t)$ and $Q_{n-1}(t)=q_n(t)/(t-\xi_k)$, we get

$$\int Q_{n-1}(t)\,t^k\,e^{i\varphi(t)}\,d\mu_1(t)=0,\qquad k=0,\dots,n-2; \tag{4}$$

but if $\xi_k\in\cup_j[a_j,b_j]$, then μ_1 is again positive and $\theta(\xi_k)=\pi$, so we are back to prove the lemma with $n-1$ instead of n, d_n-1 instead of d_n, Q_{n-1} instead of q_n, and μ_1 instead of μ. Proceeding recursively, we reach a situation where q_n has no zero on $\cup_j[a_j,b_j]$ or where $n=1$ and either its only root lies on $\cup_j[a_j,b_j]$ or else $d_n=0$. But if $q_1(t)=t-\xi_1$ with $\theta(\xi_1)=\pi$, then (3) is trivial for the left-hand side vanishes, while if $n=1$ and $d_n=0$, it is also trivial since $V(\varphi)$ must be at least π in order that the $L(d\mu)$-mean of $e^{i\varphi}$ be zero (see the argument that follows below).

Thus we can safely assume that q_n has no zero on $\cup_j[a_j,b_j]$ in which case it has a well-defined and smooth argument on a neighborhood of the latter. Then $H(t)=\arg\big(q_n(t)e^{i\varphi(t)}\big)$ is in turn well-defined of bounded variation there. The function H has left and right limit $H(x-0)$ and $H(x+0)$, respectively, at every $x\in\cup_j[a_j,b_j]$. By the *graph* of H on $\cup_j[a_j,b_j]$, we mean the set

$$\{(x,y)\,|\,x\in\cup_j[a_j,b_j],\ y\in(H(x-0),H(x)]\cup[H(x),H(x+0))\},$$

where the interval $(H(x-0),H(x)]$ (resp.$[H(x),H(x+0))$) is omitted if x is equal to some a_j (resp. some b_j). The graph can be visualized as the ordinary graph plus all the vertical segments that represent a jump of H. We will show that the respective variations $V(H,[a_j,b_j])$ of H over the intervals $[a_j,b_j]$ satisfy

$$\sum_{j=1}^{m}V(H,[a_j,b_j])\geq(n-m+1)\pi. \tag{5}$$

This is sufficient to conclude (3), because by the monotonicity of $\arg(t-\xi_k)$ we have:

$$\begin{aligned}\sum_{j=1}^{m}V(H,[a_j,b_j]) &\leq V(\varphi)+\sum_{j=1}^{m}\sum_{k=1}^{d_n}V\big(\arg(t-\xi_k),[a_j,b_j]\big)\\ &= V(\varphi)+\sum_{k=1}^{d_n}\sum_{j=1}^{m}\mathrm{Angle}(\xi_k,[a_j,b_j])=V(\varphi)+\sum_{k=1}^{d_n}\theta(\xi_k).\end{aligned}$$

Thus, it has left to show (5). Suppose to the contrary that

$$\sum_{j=1}^{m} V(H,[a_j,b_j]) < (n-m+1)\pi. \tag{6}$$

Let $N(u,H,[a_j,b_j])$ be the (possibly infinite) number of intersections of the line $y=u$ with the graph of H over $[a_j,b_j]$. In this count, the intersection at (x,y) is double if u lies in both $(H(x-0),H(x))$ and $(H(x),H(x+0))$.

By the Banach-Kestelman lemma [10]

$$\int_{-\infty}^{\infty} N(u,H,[a_j,b_j])du = V(H,[a_j,b_j]). \tag{7}$$

Thus if we let

$$N_\pi(u,H,[a_j,b_j]) = \sum_{l=-\infty}^{\infty} N(u+l\pi,H,[a_j,b_j]),$$

we get from (7) that

$$\int_0^\pi N_\pi(u,H,[a_j,b_j])du = V(H,[a_j,b_j]).$$

Adding up over $j=1,\ldots,m$, we can infer from (6) that there is a set $E\subset[0,\pi)$ of positive Lebesgue measure such that

$$\sum_{j=1}^{m} N_\pi(u,H,[a_j,b_j]) < n-m+1 \text{ whenever } u\in E. \tag{8}$$

Pick $u_0\in E$ which is not a value $H(a_j)$, $H(a_j+0)$, $H(b_j-0)$ or $H(b_j)$, nor equal to any $H(x-0)$, $H(x)$, or $H(x+0)$ for x a discontinuity point of H. This is certainly possible since such values form a countable set. We will show that there is a polynomial R_{n-1} of degree at most $n-1$ (it can be chosen with real zeros) such that either

$$\Im\left\{e^{-iu_0}q_n(t)e^{i\varphi(t)}R_{n-1}(t)\right\} \geq 0, \quad t\in\cup_j[a_j,b_j], \tag{9}$$

or

$$\Im\left\{e^{-iu_0}q_n(t)e^{i\varphi(t)}R_{n-1}(t)\right\} \leq 0, \quad t\in\cup_j[a_j,b_j], \tag{10}$$

and in any case the inequality is strict except for finitely many t's. However, (9) and (10) are then equally impossible, because the integral of

$$e^{-iu_0}q_n(t)e^{i\varphi(t)}R_{n-1}(t)$$

against $d\mu(t)$ must be zero by (1) and μ has infinite support (recall $d\lambda = e^{i\varphi(t)}d\mu(t)$). This contradiction will prove (5) pending the proof of either (9) or (10) with strict inequality at all but finitely many t's.

To construct R_{n-1}, paint red on the plane all the strips $u_0 + 2l\pi < y < u_0 + (2l+1)\pi$, $l = 0, \pm 1, \pm 2, \ldots$, and let G be their union; and paint the remaining strips $u_0 + (2l-1)\pi < y < u_0 + 2l\pi$, $l = 0, \pm 1, \pm 2, \ldots$ of the plane white. Let us also agree that the boundary ∂G of these regions (i.e. the union of the lines $y = u_0 + 2l\pi$, $l = 0, \pm 1, \pm 2, \ldots$) gets both colors. The graph of H over $[a_j, b_j]$ intersects the boundary ∂G of G at those abscissa x where either x is a continuity point of H and $H(x) \in \partial G$ or else x is a discontinuity point and at least one of the vertical segment $(H(x-0), H(x))$, $(H(x), H(x+0))$ is bicolour (one of the segments is missing if $x = a_j$ or $x = b_j$, and remember that $H(x)$ cannot lie on ∂G). In the first case the multiplicity of the intersection is 1, and in the second case it is equal to the number of bicolour segments, *i.e.* 1 or 2 (remember all the jumps are $\leq \pi$ and neither $H(x-0)$ nor $H(x)$ or $H(x+0)$ can lie on ∂G). By definition the sum of all multiplicities is $N_\pi(u_0, H, [a_j, b_j])$, which is finite because of (8); in particular there are only finitely many intersections and they occur at isolated abscissa. Among those, we further distinguish the crossing abscissa $x \in [a_j, b_j]$ where the graph of H crosses ∂G, meaning that either x is a continuity point of H and there is $\varepsilon > 0$ such that $(t, H(t))$ is red for $t \in (x-\varepsilon, x]$ (resp. $t \in [x, x+\varepsilon)$) and white for $t \in (x, x+\varepsilon)$ (resp. $t \in (x-\varepsilon, x)$, or else x is a discontinuity point of H and $(x, H(x-0))$, $(x, H(x))$, $(x, H(x+0))$ are not all of the same colour (remember none of them can lie on ∂G). It is straighforward to check with that definition that if $(x_1, H(x_1))$ lies interior to G and $(x_2, H(x_2))$ lies interior to the complement of G, some crossing has to occur between x_1 and x_2 (i.e. the intermediate value theorem holds). We also say that the graph of H starts in G (resp. outside G) at a_j if $(a_j, H(a_j))$ lies interior to G (resp. to the complement of G); here one should recall that $(a_j, H(a_j)$ cannot lie on ∂G by our choice of u_0.

We need now define arguments on the real line for functions involving a polynomial factor having real zeros. The definition of the argument at a zero is immaterial, because the function will be later integrated against μ and the contribution of the integrand at that point is zero anyway. For definiteness, we may assume for instance that the argument is left continuous. The definition of the jump of the argument across the zero also involves an arbitrary choice for the jump can be chosen to be any $m\pi + k2\pi$ where m is the multiplicity and $k \in \mathbf{Z}$. This choice is not relevant either in what follows, because only the color of the points of the graph will matter and that color is invariant under jumps in y of the form $k2\pi$. For definiteness again, let us agree that each zero produces a jump of $m\pi$ in the argument of $(x-t)$ when t crosses x from the left to the right.

If $R(t) = \prod(t_j - t)$ is a polynomial with real zeros t_j, it is straightforward to check with that definition that

$$\arg\bigl(q_n(t)e^{i\varphi(t)}R(t)\bigr) = H(t) + \sum_{t_j < t} \pi, \qquad t \neq t_j$$

where the t_j are repeated according to their multiplicity.

Now, if $x_1, \ldots, x_{n_1}$ are the crossing abscissa of H over $[a_1, b_1]$, and if we pick

$$R_1(t) = \Pi_{j=1}^{n_1}(x_j - t)^{m_j}$$

where m_j is the multiplicity of x_j, a moment's thinking will convince the reader (see the next paragraph) that the graph of $\arg\big(q_n(t)e^{i\varphi(t)}R_1(t)\big)$ over $[a_1, b_1]$ is either red or white (depending whether a_1 is a crossing point and whether the graph starts in or outside G at a_1). Furthermore, the graph lies in the interior of the red or white region except possibly at finitely many places $(x, H(x))$ where x may be an intersection abscissa which is not crossing (hence $\arg\big(q_n(x)e^{i\varphi(x)}\big) \in \partial G$) or a crossing abscissa (hence a zero of R_1) where H is discontinuous. Hence

$$\text{either} \quad \Im\big\{e^{-iu_0}q_n(t)e^{i\varphi(t)}R_1(t)\big\} > 0 \quad \text{or} \quad \Im\big\{e^{-iu_0}q_n(t)e^{i\varphi(t)}R_1(t)\big\} < 0 \tag{11}$$

holds true on $[a_1, b_1]$, except at the zeros of R_1 and at those points t where the graph of H hits ∂G without crossing; such points are intersection points (hence finite in number) at which the left-hand side of the inequalities (11) vanishes by construction.

Geometrically, the above procedure may be described as follows. Suppose to fix ideas that the graph of H starts in G. Then, a zero of R_1 placed to a crossing point where the graph would leave a red strip pushes the argument down by π, so the graph continues in the next red strip, and if the graph does not leave a red strip but only the value at the crossing point is white because of a jump (so that the crossing has multiplicity 2), we put a double zero so as to push the graph down by 2π (thus its stays red) while zeroing R_1 at the crossing point itself.

Next, we replace $H(t)$ by $\arg\big(q_n(t)e^{i\varphi(t)}R_1(t)\big)$ and we proceed in the same manner on $[a_2, b_2]$, observing that on $[a_2, b_2]$ the factor R_1 shifts H by a multiple of π thus it does not change the crossing points nor their multiplicity. We thus let

$$R_2 = \Pi_{j=n_1+1}^{n_2}(x_j - t)^{m_j}$$

where $x_{n_1+1}, \ldots, x_{n_2}$ are the crossing abscissa of the graph of H over $[a_2, b_2]$ with multiplicities m_j, and we find that

$$\begin{aligned} \text{either} \quad \Im\big\{e^{-iu_0}q_n(t)e^{i\varphi(t)}R_1(t)R_2(t)\big\} \quad &> \quad 0 \\ \text{or} \quad \Im\big\{e^{-iu_0}q_n(t)e^{i\varphi(t)}R_1(t)R_2(t)\big\} \quad &< \quad 0 \end{aligned} \tag{12}$$

except perhaps at finitely many places where the left-hand side vanishes. Note that the factor $R_2(t)$ does not influence the argument on $[a_1, b_1]$, so we preserve for $\Im\big\{e^{-iu_0}q_n(t)e^{i\varphi(t)}R_1(t)R_2(t)\big\}$ whichever inequality was prevailing in (11). In case the latter is not the same as in (12), we put an extra zero to R_2 at a_2 so that the whole graph of $\arg(e^{-iu_0}q_n(t)e^{i\varphi(t)}R_1(t)R_2(t))$ over $[a_2, b_2]$ gets shifted by π and the valid inequality in (12) gets swaped.

Continuing in this fashion, we find $R = R_1R_2 \ldots R_m$ such that either (9) or (10) holds with R in place of R_{n-1}, and the inequality is strict except perhaps at finitely many places where the left-hand side vanishes.

To check the degree of R, we observe that we needed to proceed as many zeros as the total number of crossings, which is

$$\sum_{j=1}^{m} N_\pi(u_0, H, [a_j, b_j]),$$

plus possibly one extra zero to each a_j with $j \geq 2$ which is at most $m-1$. Thus $R = R_{n-1}$ has degree at most $n-1$ by (8), as was to be shown.

∎

To establish Theorem 3.1, we need another lemma, that is reminiscent of the previous proof, but is actually much simpler because we do not control the degree of the polynomial involved.

Lemma 3.4 *If φ is of bounded variation on an interval $[a,b]$, then there is a polynomial T and a $\delta > 0$ such that*

$$\arg\bigl(e^{i\varphi(t)}T(t)\bigr) \in \left[-\frac{\pi}{2}+\delta, \frac{\pi}{2}-\delta\right], \qquad t \in [a,b], \quad T(t) \neq 0. \tag{13}$$

Proof. Up to the addition of a piecewise constant function taking values in multiples of 2π, we may assume as in the proof of Lemma 3.2 that the jumps of φ all lie in the interval $(-\pi, \pi]$.

We suppose first that φ is left-continuous. Let $t_1, \ldots, t_s$ denote the points where $\varphi(t_j + 0) - \varphi(t_j) = \pi$, and set

$$\psi(t) = \varphi(t) - \pi \sum_{j=1}^{s} \chi_{(t_j, b]}(t),$$

where χ_E indicates the characteristic function of the set E. The function ψ is of bounded variation on $[a,b]$, it is left continuous, and all its jumps are strictly less than π in absolute value. Let

$$\gamma = \sup_t |\psi(t+0) - \psi(t)|$$

be the supremum of the absolute value of its jumps. Then $\gamma < \pi$, and if $\gamma/2 < \gamma_1 < \pi/2$, then there is a continuous function g with $|\psi - g| < \gamma_1$. Let $\delta < (\pi/2 - \gamma_1)/2$, and T_0 be a polynomial such that

$$|e^{-ig(t)} - T_0(t)| < \frac{\delta}{2}, \qquad t \in [a,b].$$

Shrinking δ if necessary ensures that T_0 has no zero on $[a,b]$. Now, $|1 - e^{ig(t)}T_0(t)| < \delta/2$ hence

$$|e^{i(\psi(t)-g(t))} - e^{i\psi(t)}T_0(t)| < \delta/2,$$

implying that

$$\Bigl|(\psi(t) - g(t)) - \arg\bigl(e^{i\psi(t)}T_0(t)\bigr)\Bigr| < \delta.$$

Consequently

$$\arg\bigl(e^{i\psi(t)}T_0(t)\bigr) \in \left[-\frac{\pi}{2}+\delta, \frac{\pi}{2}-\delta\right]$$

because $\psi(t) - g(t) \in [-\pi/2 + 2\delta, \pi/2 - 2\delta]$. Noticing that

$$\arg\big(e^{i\psi(t)}T_0(t)\big) = \arg\Big(e^{i\varphi(t)}T_0(t)\prod_{j=1}^{s}(t_j - t)\Big), \quad t \neq t_j,$$

the lemma follows with $T(t) = T_0(t)\prod_{j=1}^{s}(t_j - t)$.

In the general case, φ can be written as the sum of a left-continuous function of bounded variation, say φ_1, plus an absolutely convergent series of functions $\sum_m h_m$ where $h_m = 0$ except at some point x_m where, say $h_m(x_m) = y_m$. We apply the previous part of the proof to φ_1, and we select finitely many y_m such that the sum of the moduli of the remaining ones is less that $\delta/2$. Putting an additional double zero of T at each x_m corresponding to a selected y_m yields (13) with δ replaced by $\delta/2$.

■

Proof of Theorem 3.1. Let ν be a weak* limit point of $\{\nu_n\}$. By Corollary 3.3 this ν is supported on S and it has total mass 1. Following the proof in the preceding section, it is again sufficient to show that there is a constant D such that the potential U^ν of ν equals D quasi-everywhere on S.

Let us suppose first that the zeros of q_n remain in a bounded set. We argue by contradiction exactly as in Theorem 2.1, following the argument there up to the choice of P_{n-2}. In this connection, it must be stressed that the principle of descent that led to (2) is valid because we assume that no zero of q_n can go to infinity. In fact, up to (5) everything remains true word for word with the degree n replaced by $d_n = \deg(q_n)$. In particular, (5) takes the form

$$\int_{[x_n-\rho/2,x_n+\rho/2]} |q_n|^2 d\mu \geq e^{-2d_n d + d_n \tau}. \tag{14}$$

Now, we fix a polynomial $T(x)$, say of degree k, as in Lemma 3.4, where in that lemma $[a, b]$ is some interval containing S. If c_k is the leading coefficient of T, we get from the Boutroux-Cartan lemma [13] that

$$|T(t)| \geq |c_k| \left(\frac{\varepsilon}{8e^{n\lambda(\varepsilon)+1}}\right)^k$$

outside a union of at most k open disks, the sum of whose radii is at most $\varepsilon/(4e^{n\lambda(\varepsilon)})$. Using this it is easy to see from the proof of (5) that (14) is also true in the form

$$\int_{[x_n-\rho/2,x_n+\rho/2]} |q_n|^2 |T(x)| d\mu \geq e^{-2d_n d + d_n \tau} \tag{15}$$

when d_n is large enough, provided $\varepsilon > 0$ is so small that $\lambda(\varepsilon)(k + L) < \tau$.

Now, for small $0 < \eta < \mathrm{diam}(S)/2$, let $\alpha_{n,1}, \alpha_{n,2}, \ldots, \alpha_{n,k+1}$ be $k + 1$ zeros of q_n in the the η-neighborhood of the interval $[y_0 - \rho/4, y_0 + \rho/4]$ (recall (2) for the definition of y_0).

For sufficiently large n (hence large d_n since $n - d_n$ is bounded as we saw after the proof of Lemma 3.2), there are certainly $k+1$ zeros lying in that neighborhood, because y_0 was a point in the support of the limit measure ν. Consider the polynomial

$$P_{n-1}(t) = \frac{\overline{q_n(t)}T(t)}{\prod_{j=1}^{k+1}(t - \overline{\alpha_{n,j}})}$$

of degree at most $n-1$. By choosing η sufficiently small, we can make the absolute value of the argument of

$$1\Big/\prod_{j=1}^{k+1}(t - \overline{\alpha_{n,j}})$$

smaller than $\delta/2$ for all $t \in \mathbf{R} \setminus [y_0 - \rho/2, y_0 + \rho/2]$. Since for all $t \in S$, whith $T(t) \neq 0$, the argument of $q_n(t)\overline{q_n(t)}e^{i\varphi(t)}T(t)$ is in $(-\pi/2 - \delta, \pi/2 - \delta)$, it follows that the argument of $q_n(t)P_{n-1}(t)e^{i\varphi(t)}$ lies in $(-\pi/2 + \delta/2, \pi/2 - \delta/2)$ for each $t \in S \setminus [y_0 - \rho/2, y_0 + \rho/2]$ which is not a zero of P_{n-1}, and so we get from (15)

$$\begin{aligned}
\Re \int_{S\setminus[y_0-\rho/2, y_0+\rho/2]} q_n(t)P_{n-1}(t)d\lambda(t) &\geq \Re \int_{[x_n-\rho/2, x_n+\rho/2]} q_n(t)P_{n-1}(t)d\lambda(t) \\
&\geq \frac{\sin(\delta/2)}{(2\mathrm{diam}(S))^{k+1}} \int_{[x_n-\rho/2, x_n+\rho/2]} |q_n|^2|T(x)|d\mu \\
&\geq \frac{\sin(\delta/2)}{(2\mathrm{diam}(S))^{k+1}} e^{-2d_n d + d_n \tau} \qquad (16)
\end{aligned}$$

as soon as $n \in \mathbf{N}_2$ is large enough. On the other hand, the limit distribution of the zeros of the polynomials P_{n-1}, $n \in \mathbf{N}_2$ is again ν, because T is fixed and we only discarded a fixed amount of $k+1$ zeros $\overline{\alpha}_{n,1}, \overline{\alpha}_{n,2}, \ldots, \overline{\alpha}_{n,k+1}$ from $\overline{q}_n$, these being in a fixed compact set, while the asymptotic zero distribution of $\overline{q}_n$ and q_n is the same. Therefore, as we already remarked after (2), we also have:

$$|P_{n-1}(y)| \leq |c_k|e^{-d_n d}, \qquad y \in [y_0 - \rho, y_0 + \rho]$$

for sufficiently large n in a suitable subsequence. Thus we obtain from this and (2) the estimate

$$\left| \int_{y_0-\rho/2}^{y_0+\rho/2} q_n(t)P_{n-1}(t)d\lambda(t) \right| \leq Ce^{-2d_n d}. \qquad (17)$$

But the sum of the two integrals on left of (16) and (17) should be zero by orthogonality, which, in view of the estimates in these formulae, is clearly not the case. This contradiction proves the theorem when the zeros of q_n remain bounded.

In the general case, we see from Corollary 3.3 that at most a fixed number of zeros of q_n, say N, can leave every compact subset of $\mathbf{C}$ as n goes large. Denote the zeros that actually tend to infinity through a subsequence of the n's with $n \in \mathbf{N}_1$, say for $n \in \mathcal{N}$, by

$$\xi_1^{(n)}, \xi_2^{(n)}, \ldots, \xi_m^{(n)},$$

where m is a fixed integer which is no larger than N, and put

$$p_m^{(n)}(t) = \Pi_{j=1}^m \left(t - \xi_j^{(n)}\right), \qquad q_n^*(t) = \frac{q_n(t)}{p_m^{(n)}(t)}.$$

Note that q_n^* has degree at most $n-m$, that its zeros remain bounded, and that it satisfies the orthogonality relations:

$$\int q_n^*(t)\, t^k\, e^{i\varphi(t)} \left|p_m^{(n)}(t)\right|^2 d\mu(t) = 0, \qquad k = 0, \ldots, n-m-1. \tag{18}$$

Furthermore, the degree of q_n^* is at most $n-m$, so in what follows $n-m$ plays the role of n from the first part of the proof. For n large enough we certainly have that $\xi_j^{(n)} \neq 0$, so we can define a sequence of positive measures by setting

$$d\mu^{(n)}(t) = \left|\frac{p_m^{(n)}(t)}{\Pi_{j=1}^m \xi_j^{(n)}}\right|^2 d\mu(t), \qquad n \in \mathcal{N}$$

and upon renormalizing (18) we get:

$$\int q_n^*(t)\, t^k\, e^{i\varphi(t)}\, d\mu^{(n)}(t) = 0, \qquad k = 0, \ldots, n-m-1.$$

Because the asymptotic zero distribution of q_n and q_n^* is the same, we would be done if only we could apply the first part of the proof to q_n^* with $\mu^{(n)}$ instead of μ. Although $\mu^{(n)}$ depends on n, this is indeed possible provided that the constant c in (1) can be made uniform with respect to n, and provided also that the total mass remains bounded independently of n, for these are the only facts that were used on μ beyond positivity. But it is trivial that $\mu^{(n)}$ meets these requirements because

$$\frac{\left|p_m^{(n)}(t)\right|^2}{\Pi_{j=1}^m |\xi_j^{(n)}|^2} \longrightarrow 1, \qquad n \to \infty,\ n \in \mathcal{N}$$

uniformly on S as n goes to infinity, and the proof of Theorem 3.1 is now complete.

■

4 Rational approximants to Markov functions

Let μ be a positive measure with infinite compact support $S \subset (-1, 1)$, and consider the associated Markov function

$$M(z) = \int \frac{1}{z-t} d\mu(t). \tag{1}$$

If we form the diagonal Padé approximants to $M(z)$ at ∞, we mentioned already that the denominators are just the orthogonal polynomials with respect to μ. If however we form a best rational L^2 approximant of degree at most n to $M(z)$ on the unit circle $\mathbf{T}$, say p_{n-1}/q_n (by Parseval's theorem it must vanish at infinity) then its denominator

$$q_n(t) = \prod_{j=1}^{n}(t-\alpha_j)$$

satisfies the orthogonality conditions:

$$\int \frac{q_n(t)}{Q_n^2(t)} t^k d\mu = 0, \qquad 0 \le k < n, \tag{2}$$

where

$$Q_n(t) = \prod_{j=1}^{n}(t - 1/\overline{\alpha_j})$$

is the polynomial whose zeros are reflected from those of q_n across the unit circle (if $\alpha_j = 0$, then the corresponding factor is missing from Q_n). In fact, that q_n must be real is not immediate but true [1], and then (2) follows from [7] (which deals with real approximants only). In particular q_n is the n-th orthogonal polynomial associated to the positive measure $Q_n^{-2}(t)d\mu(t)$.

However, a best approximant of degree n need not be unique, hence (2) may have several solutions [4]. Moreover, equation (2) is met not only by the denominator of a best approximant, but more generally by the denominator of each critical point of the L^2-error in degree n, including local minima, saddles and so on. Sufficient condition for uniqueness can be found in [7, 6], but the theorem we shall prove is valid for any sequence of polynomials satisfying (2). Therefore we simply assume that we are given such a sequence, which is consistently denoted by $\{q_n\}$. From Lemma 5.2 to come (put $m = 1$ in that lemma) q_n necessarily has real coefficients so that, being the n-th orthogonal polynomial associated to the positive measure $Q_n^{-2}(t)d\mu(t)$, it has n simple zeros lying in the convex hull of S.

We are interested in the limit distribution of the zeros of q_n, so we let ν_n be their normalized counting measure.

About μ we assume exactly as in Section 2 that S is regular and that (1) holds.

To formulate our result we need introduce the Green equilibrium measure of S with respect to the unit disk. Let

$$g(z,a) = \log\left|\frac{1-\overline{a}z}{z-a}\right|$$

be the Green function of the unit disk with pole at a. To each probability measure σ with support in S we associate the Green energy

$$\int\int g(z,a)d\sigma(z)d\sigma(a).$$

Now, among all probability measures supported on S, there is one and only one measure Ω_S minimizing the Green energy, which is called the Green equlibrium measure of S associated with the unit disk (see e.g. [15]).

Theorem 4.1 *Suppose that the support S of μ is regular with respect to the Dirichlet problem in $\mathbf{C} \setminus S$, and that (1) holds. Then ν_n tends to the Green equilibrium measure Ω_S of S in the weak* topology as n tends to ∞.*

Proof. We start the proof with the observation that the problem is invariant under Möbius transformations $z \to (z-a)/(1-az)$, $a \in (-1,1)$. It is known that the Green equilibrium measure is invariant under Möbius transformation, and we claim that the statement is also invariant. In fact, let $w = (z-a)/(1-az)$ be a new complex variable, reserving the special notation $\tau = (t-a)/(1-at)$ when $t \in (-1,1)$, and define a new measure ν by setting:

$$\frac{d\nu(\tau)}{1+a\tau} = d\mu(t).$$

Observe that the support of ν, being the image of S under the conformal map $z \mapsto w$, is again a regular compact subset of $(-1,1)$, and that (1) will hold for ν (with a different c).

Note also that for any bounded measurable function h on $(-1,1)$ we have:

$$\int h\left(\frac{t-a}{1-at}\right) d\mu(t) = \int h(\tau)\,\frac{d\nu(\tau)}{1+a\tau}.$$

If λ_j is the image of α_j under the mapping $z \mapsto w$, then $1/\overline{\lambda_j}$ is the image of $1/\overline{\alpha_j}$. Therefore, the orthogonality relations

$$\int t^k \left(\prod_1^n \frac{t-\alpha_j}{(t-1/\overline{\alpha_j})^2}\right) d\mu(t) = 0, \qquad k = 0,1,\dots n-1$$

take (modulo a multiplicative constant) the form:

$$\int \left(\frac{\tau+a}{1+a\tau}\right)^k \left(\prod_1^n \frac{(\tau-\lambda_j)(1+a\tau)}{(\tau-1/\overline{\lambda_j})^2}\right) \frac{d\nu(\tau)}{1+a\tau} = 0, \qquad k = 0,1,\dots n-1.$$

Here

$$\left(\frac{\tau+a}{1+a\tau}\right)^k (1+a\tau)^{n-1}, \qquad k = 0,1,\dots,n-1$$

generate all polynomials in τ of degree at most $n-1$, hence the preceding relations are the same as the orthogonality relations:

$$\int \tau^k \left(\prod_1^n \frac{\tau-\lambda_j}{(t-\overline{\lambda_j})^2}\right) d\nu(\tau) = 0, \qquad k = 0,1,\dots n-1. \tag{3}$$

If, on the basis of (3), we are now able to prove the theorem for ν, it will hold for μ as well by invariance of the Green equilibrium measure under Möbius transforms, as announced.

Back to our problem, having seen that it is invariant under Möbius transformations $z \mapsto (z-a)/(1-az)$ where $a \in (-1,1)$ can be chosen at will, we pick it to the left of S and make this preliminary transformation so as to be able to assume in what follows that $S \subset (0,1)$.

After these we prove the promised asymptotic behavior of the zeros, closely following the proof in Section 2. Let ν be a weak* limit point of $\{\nu_n\}$, say $\nu_n \to \nu$ as $n \to \infty$, $n \in \mathbf{N}_1$. Since, we saw at the beginning of this section, q_n is the n-th orthogonal polynomial with respect to the positive measure $d\mu(t)/Q_n^2(t)$, each subinterval of $\mathbf{R} \setminus S$ can contain only one-one zero of q_n. Therefore, ν is a probability measure supported on S.

Together with ν_n we also consider the normalized zero counting measure of Q_n, let it be σ_n. It is immediate that if $\nu_n \to \nu$, then $\sigma_n \to \sigma$, where σ is the reflection of ν across the unit circle. In particular, σ is supported outside the unit disk on the compact subset S^{-1} of $\mathbf{R}$ (remember that $S \subset (0,1)$, and its potential is continuous in the unit disk. Likewise, taking into account that the zeros of Q_n remain in the convex hull of S^{-1}, it is easy to see that the potentials U^{σ_n} (i.e. $\log|Q_n|^{-1/n}$) are uniformly equicontinuous and bounded on S. In particular, up to refining the subsequence $\mathbf{N}_1$, we may assume that U^{σ_n} converges uniformly on S and the limit is necessarily U^σ because for $t \in S$ $z \to \log|z-t|$ is continuous in a neighborhood of S^{-1}.

We claim that it is enough to prove there is a constant D such that the potential $U^{\nu-\sigma}$ of $\nu - \sigma$ equals D quasi-everywhere on S. In fact, then the lower semi-continuity of $U^{\nu-\sigma}$ implies that $U^{\nu-\sigma}(x) \leq D$ for all $x \in S$. Thus, ν has finite logarithmic energy. Let $\widehat{\sigma}$ be the balayage of the measure σ onto S. The regularity of S with respect to the Dirichlet problem implies that $U^{\widehat{\sigma}}(x) = U^\sigma(x) + c_1$ for some constant c_1 and all $x \in S$. In particular $\widehat{\sigma}$ has finite logarithmic energy, and the potential $U^{\nu-\widehat{\sigma}}(x)$ is equal to $D - c_1$ for quasi-every $x \in S$. Now, integrating the equality "$U^{\nu-\widehat{\sigma}}(x) = D - c_1$ for quasi-every $x \in S$" against $\nu - \widehat{\sigma}$ we get from the fact that $(\nu - \widehat{\sigma})(S) = 0$ the equality:

$$0 = \int U^{\nu-\widehat{\sigma}} d(\nu - \widehat{\sigma}) = \int\int \log\frac{1}{|z-t|} d(\nu-\widehat{\sigma})(t) d(\nu - \widehat{\sigma})(z),$$

and it is known for a signed measure with finite logarithmic energy and of total mass zero that the logarithmic energy can be zero only if the measure is zero (see [15, Lemma I.1.8]). Thus, $\nu = \widehat{\sigma}$. Let now $\tilde{\sigma}$ and $\tilde{\nu}$ denote the respective balayages of σ and ν onto the unit circle. Seeing that the balayage of the Dirac delta δ_α at α onto the unit circle is the same as the balayage of the Dirac delta $\delta_{1/\overline{\alpha}}$ at the reflected point $1/\overline{\alpha}$, it follows that $\tilde{\sigma} = \tilde{\nu}$. However, forming the balayage of σ onto S can be done in two steps: first form the balayage onto the unit circle, then form the balayage of the so obtained measure $\tilde{\sigma}$ onto S. Thus, using as before a "hat" to denote the balayage onto S and a "tilde" to denote the balayage on the unit circle, we get that $\widehat{\sigma} = \widehat{\tilde{\sigma}}$, and so we can write

$$\nu = \widehat{\sigma} = \widehat{\tilde{\sigma}} = \widehat{\tilde{\nu}},$$

i.e. ν has the property that if we form its balayage onto the unit circle and then form the balayage of that measure back onto S, we obtain ν again. This, however, characterizes the Green equilibrium measure Ω_S (see [15, Theorem VIII.2.6]), hence $\nu = \Omega_S$. Since this is true for all weak* limit points of $\{\nu_n\}$, the whole sequence converges to Ω_S, *as claimed.*

Thus, it has left to prove there is a constant D such that the potential of $\nu - \sigma$ equals D quasi-everywhere on S. We closely follow the reasoning in Theorem 2.1, but there are some minor changes that we have to indicate.

Suppose to the contrary that the claim is not true, and there are d, $\tau > 0$ and two sets $E_- \subset S$ and $E_+ \subset S$ of positive capacity such that

$$U^{\nu-\sigma}(x) \le d - 2\tau \quad \text{for } x \in E_- \qquad \text{and} \qquad U^{\nu-\sigma}(x) \ge d + \tau \quad \text{for } x \in E_+$$

(note that $U^{\nu-\sigma}$ is certainly finite quasi-everywhere on S since both U^ν and U^σ are). We claim that there is $y_0 \in \mathrm{supp}(\nu)$ such that $U^{\nu-\sigma}(y_0) > d$. In fact, in the opposite case $U^\nu(x) \le U^\sigma(x) + d$ for all $x \in S$, which implies first of all that ν has finite logarithmic energy and then by the principle of domination that the same inequality holds for all $z \in \mathbf{C}$, which is a contradition, for on E_+ we have bigger values for $U^{\nu-\sigma}$.

According to the principle of descent (remember that the support of ν_n remains in the convex hull of S and that U^{σ_n} converges uniformly to U^σ on the disk), we have that

$$\liminf_{n\to\infty,\ n\in\mathbf{N}_1} U^{\nu_n-\sigma_n}(y_n) \ge U^{\nu-\sigma}(y_0) > d$$

for any sequence $y_n \to y_0$. Therefore, there is $\rho > 0$ and N_1 such that, for $y \in [y_0 - \rho, y_0 + \rho]$ and $n \ge N_1$, $n \in \mathbf{N}_1$, the inequality $U^{\nu_n-\sigma_n}(y) \ge d$ holds which is the same as

$$|q_n(y)/Q_n(y)| \le e^{-nd}, \qquad y \in [y_0 - \rho, y_0 + \rho]. \tag{4}$$

As before, this inequality remains true (for sufficiently large n) if we replace q_n by some monic polynomial q_n^* with $\deg(q_n^*) = (1 + o(1))n$, provided that the zeros of q_n^* remain in a compact set of the unit disk and their asymptotic distribution is still ν.

In another connection, the lower envelope theorem implies that, for quasi-every $x \in E_-$, we have

$$\liminf_{n\to\infty,\ n\in\mathbf{N}_1} U^{\nu_n-\sigma_n}(x) = U^{\nu-\sigma}(x) \le d - 2\tau,$$

hence for some subsequence $\mathbf{N}_2 \subset \mathbf{N}_1$ and sufficiently large $n \in \mathbf{N}_2$, we get:

$$M_n := \max_{x\in S} |q_n(x)/Q_n(x)| \ge e^{-n(d-\tau)}. \tag{5}$$

We let x_n be a point where the maximum is attained.

As in the proof of Theorem 2.1, we now need an estimate for q_n away from S in terms of M_n. To this effect, we set $w(x) = 1/|Q_n(x)|^{1/n}$, $x \in S$, and we consider the weighted energy problem for w on S (see [15, Theorem I.1.3]). If we denote by $\widehat{\sigma_n}$ the balayage of σ_n onto S, we deduce from the definition of balayage and the regularity of S that $w(x) = \exp(U^{\widehat{\sigma_n}}(x) - c_n)$

for every $x \in S$ and some constant c_n. Then it follows from [15, Theorem I.3.3] that the equilibrium measure associated with this weighted energy problem is precisely $\widehat{\sigma_n}$. So, from the generalized Bernstein–Walsh lemma [15, Theorem III.2.1], we obtain:

$$|q_n(z)| \leq M_n \exp(-nU^{\widehat{\sigma_n}}(z) + nc_n). \tag{6}$$

Note that for $z = x_n$ the exponential factor on the right is just $|Q_n(x_n)|$. Since $U^{\widehat{\sigma_n}}(z)\Big|_S$ is continuous on S, it follows from the continuity theorem for logarithmic potentials (Maria's theorem) [15, Theorem II.3.5] that $U^{\widehat{\sigma_n}}$ is continuous on the whole plane, and actually the continuity on S is uniform in n (this follows from the equicontinuity of the potentials U^{σ_n} on S, and from the very proof of the continuity theorem). Hence, for every $\theta > 0$, there is $\varepsilon > 0$ such that for $|z - x_n| \leq 2\varepsilon$ we have

$$|U^{\widehat{\sigma_n}}(z) - U^{\widehat{\sigma_n}}(x_n)| \leq \theta.$$

All these imply in view of (6) that, for $|z - x_n| \leq 2\varepsilon$,

$$|q_n(z)| \leq |Q_n(x_n)| M_n e^{n\theta}.$$

Then, by Cauchy's formula,

$$|q_n'(z)| \leq \frac{|Q_n(x_n)| M_n e^{n\theta}}{\varepsilon}$$

for $|z - x_n| \leq \varepsilon$ and, as in the proof of Theorem 2.1, we obtain that

$$|q_n(z)| \geq |Q_n(x_n)| M_n / 2 \tag{7}$$

for $|z - x_n| \leq \varepsilon / 2e^{n\theta}$.

From the elementary identity:

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e,$$

it easily follows that for $|z - x_n| = o(1/n)$ we have $Q_n(z) \leq 2Q_n(x_n)$ for sufficiently large n. On dividing (7) by the latter inequality, we see that

$$|q_n(z)/Q_n(z)| \geq M_n/4, \qquad |z - x_n| \leq \varepsilon/2e^{n\theta}$$

when $n \in \mathbf{N}_2$ is large enough.

From here, (1), and (5), we deduce as in the proof of Theorem 2.1 that

$$\begin{aligned} \int_{[x_n-\rho/2, x_n+\rho/2]} \frac{|q_n|^2}{|Q_n|^2} d\mu \quad &\geq \quad \left(\frac{M_n}{4}\right)^2 c \left(\frac{\varepsilon}{2e^{n\theta}}\right)^L \geq \frac{c}{16} \left(\frac{\varepsilon}{2e^{n\theta}}\right)^L e^{-2n(d-\tau)} \\ &\geq \quad e^{-2nd + n\tau} \end{aligned} \tag{8}$$

for sufficiently large $n \in \mathbf{N}_2$, say for $n \geq N_2$, provided $\theta > 0$ is so small that $\theta L < \tau$.

Now, for $n \in \mathbf{N}_2$, $n \geq \max(N_1, N_2)$ let us choose two zeros $\alpha_{n,1} < \alpha_{n,2}$ of q_n in the interval $[y_0 - \rho/2, y_0 + \rho/2]$, and consider the polynomial $P_{n-2}(t) = q_n(t)/(t - \alpha_{n,1})(t - \alpha_{n,2})$ of degree at most $n-2$ (we can find such zeros because y_0 is in the support of ν). As in the proof of Theorem 2.1 we get on the one hand from the definition of P_{n-2} and (8) that

$$\int_{t \notin (\alpha_{n,1}, \alpha_{n,2})} \frac{q_n(t)}{Q_n^2(t)} P_{n-2}(t) d\mu(t) \geq \frac{1}{(\mathrm{diam}(S))^2} e^{-2nd + n\tau}. \tag{9}$$

On the other hand, the limit distribution of the zeros of the polynomials P_{n-2}, $n \in \mathbf{N}_2$, is again ν, and these zeros remain in the convex hull of S, hence as we have already remarked after (4) we also have

$$|P_{n-2}(y)/Q_n(y)| \leq e^{-nd}, \qquad y \in [y_0 - \rho, y_0 + \rho]$$

for sufficiently large n. This and (4) together lead to the estimate:

$$\left| \int_{t \in (\alpha_{n,1}, \alpha_{n,2})} \frac{q_n(t)}{Q_n^2(t)} P_{n-2}(t) d\mu(t) \right| \leq C e^{-2nd}. \tag{10}$$

Now clearly, the sum of the two integrals on left of (9) and (10) cannot be zero. But this contradicts the orthogonality relation (2), and this contradiction proves the theorem.

■

5 Rational approximants to Cauchy transforms

In this section, we let λ be a complex measure with infinite compact support $S \subset (-1, 1)$, and we consider its Cauchy transform

$$M(z) = \int \frac{1}{z - t} d\lambda(t). \tag{1}$$

When λ is positive, then this is just the Markov function discussed in the preceding section. If we form again a best-L^2 rational approximant of degree at most n to $M(z)$ on the unit circle, say p_{n-1}/q_n, then the denominators

$$q_n(t) = \prod_{j=1}^{n} (t - \alpha_j)$$

have exact degree n, their zeros lie in the open unit disk, and they satisfy the orthogonality relations [3]:

$$\int \frac{q_n(t)}{Q_n^2(t)} t^k \, d\lambda(t) = 0, \qquad 0 \leq k < n, \tag{2}$$

where

$$Q_n(t) = \prod_{j=1}^{n}(t - 1/\overline{\alpha_j})$$

is the polynomial that has zeros at the reflected zeros of q_n across the unit circle (if $\alpha_j = 0$, then the corresponding factor is missing from Q_n). We emphasize that in the present case the zeros of q_n need not be real and so, besides λ that is in general complex-valued, the factor $q_n(t)/Q_n^2(t)$ in the orthogonality relation is also complex-valued.

Of course, even less than in the real case does (2) characterize q_n uniquely in general. There may be many solutions of degree n with zeros inside the disk, induced by several best approximants, local minima, saddles and so on. There may also be solutions of lower degree, but these have no special interpretation with respect to approximation and we shall not consider them. It is not known whether there may be solutions having zeros outside the disk.

For our purposes, we shall simply assume that we are given a sequence of solutions $\{q_n\}$ of exact degree n whose zeros lie in the unit disk. In particular, $\{q_n\}$ can be a sequence of denominators of best-$L^2(\mathbf{T})$ rational approximants to M.

Again, we are interested in the limit distribution of the zeros of q_n, so we let ν_n be their normalized counting measure.

About λ and its support S we assume the same conditions as in Section 3, namely that S is a regular set with respect to Dirichlet's problem in $\overline{C} \setminus S$, and that λ can be written as

$$d\lambda(t) = e^{i\varphi(t)}d\mu(t)$$

where φ is of bounded variation and μ satisfies the density relation (1).

Theorem 5.1 *With the preceding assumptions ν_n tends to the Green equilibrium measure Ω_S of S in the weak* topology as n tends to ∞.*

Proof. By now it should be clear how the proof proceeds: in fact Theorem 5.1 is related to Theorem 4.1 exactly as Theorem 3.1 is related to Theorem 2.1. We follow the proof of Theorems 4.1 and 3.1.

First we apply a preliminary Möbius transformation $z \mapsto (z-a)/(1-az)$, $a \in (-1,1)$ so as to ensure that $S \subset (0,1)$ (see the proof of Theorem 4.1).

Next we consider any collection of m intervals $[a_j, b_j]$ such that

$$S \subseteq \cup_{j=1}^{m}[a_j, b_j] \subset (0,1) \quad \text{with} \quad a_1 < b_1 < a_2 < \cdots < a_m < b_m.$$

Recalling the notation $\text{Angle}(\xi, [a,b]) = |\arg(a-\xi) - \arg(b-\xi)|$ for the angle in which an interval $[a,b]$ is seen from ξ, we set as in Lemma 3.2

$$\theta(\xi) = \left(\sum_{j=1}^{m} \text{Angle}(\xi, [a_j, b_j])\right),$$

which is the the total angle in which $\cup_j[a_j, b_j])$ is seen from ξ.

We shall need the analog to that lemma:

Lemma 5.2 *Let $q_n(z) = \prod_{k=1}^{n}(z - \xi_k)$ be the n-th orthogonal polynomial in the sense of* (2). *Then*

$$\sum_{k=1}^{n}(\pi - \theta(\xi_k)) \leq V(\varphi) + (m-1)\pi. \tag{3}$$

Corollary 5.3 *Under the assumptions of Theorem 5.1, for every neighborhood U of S there is a constant K_U such that q_n has at most K_U zeros outside U.*

This corollary immediately follows from Lemma 5.2 exactly as we deduced Corollary 3.3 from Lemma 3.2.

We will have to take care of the argument of $1/Q_n^2(t)$, as well. This will be done via the next lemma.

Lemma 5.4 *Let $S \subset [a, b] \subset (0, 1)$ be an interval containing the support S of λ. To every $\delta > 0$ there exists an integer l such that, for all each n large enough, there is a polynomial $T_{l,n}$ of degree at most l satisfying:*

$$\left|\frac{Q_n(t)}{|Q_n(t)|} - T_{l,n}(t)\right| < \delta, \qquad t \in [a, b]. \tag{4}$$

In particular, the argument of the ratio

$$T_{l,n}(t)/Q_n(t)$$

lies in the interval $(-2\delta, 2\delta)$ when n is large enough.

Taking these lemmas for granted, we complete the proof of Theorem 5.1 and return to the lemmas afterwards.

Fix a polynomial T of degree, say k such that, for some $0 < \delta < 1/2$ we have $\arg(e^{i\varphi(t)}T(t)) \in [-\pi/2 + 6\delta, \pi/2 - 6\delta]$ provided that $T(t) \neq 0$ (see Lemma 3.4).

Let ν be a weak* limit point of $\{\nu_n\}$, say $\nu_n \to \nu$ as $n \to \infty$, $n \in \mathbf{N}_1$. By Corollary 5.3 this ν is supported on S. Together with ν_n we also consider σ_n, the normalized zero counting measure of Q_n. By Lemma 5.4 there is a polynomial $T_{l,n}$ of degree at most some fixed number l such that (4) holds with $\delta/2$ instead of δ for all t lying in the convex hull $[a, b]$ of S. Since the sequence $T_{l,n}$ is uniformly bounded on $[a, b]$ and its degree is at most l, the coefficients of $T_{l,n}$ are in turn bounded and up to refining $\mathbf{N}_1$ we may assume that $T_{n,l}$ converges to some T_l of degree at most l. But for n large enough (4) will hold with T_l instead of $T_{l,n}$ and consequently $\arg(T_l^2(t)/Q_n^2(t)) \in [-2\delta, 2\delta]$ and $1/2 < |T_l(t)| < 2$ will hold for large $n \in \mathbf{N}_1$.

Exactly as in the proof of Theorem 4.1 it is enough to verify that there is a constant D such that the potential $U^{\nu-\sigma}$ of $\nu - \sigma$ equals D quasi-everywhere on S. *We first assume that* *no zero of Q_n goes to ∞, or equivalently that no zero of q_n goes to zero.* Then, if there is no constant D as above, we obtain as in the preceding section that (4) holds for large $n \in \mathbf{N}_1$, and then that

$$\int_{[x_n-\rho/2,x_n+\rho/2]} \frac{|q_n(t)|^2}{|Q_n(t)|^2} d\mu(t) \geq e^{-2nd+n\tau}$$

for large $n \in \mathbf{N}_2 \subset \mathbf{N}_1$ (cf. (8)). This time we actually rather need the estimate

$$\int_{[x_n-\rho/2,x_n+\rho/2]} \frac{|q_n(t)|^2}{|Q_n(t)|^2} |T(t)||T_l^2(t)| d\mu(t) \geq e^{-2nd+n\tau}, \tag{5}$$

which can be obtained by the same method, appealing to the Boutroux-Cartan lemma for T as we did to obtain (15) and taking into account that $|T_l(t)| > 1/2$.

Now, for $0 < \eta < \min\{\mathrm{diam}(S)/2, \rho/4\}$, let $\alpha_{n,1}, \alpha_{n,2}, \ldots, \alpha_{n,k+2l+1}$ be $k+2l+1$ zeros of $\overline{q_n(t)}$ in the the η-neighborhood of the interval $[y_0 - \rho/4, y_0 + \rho/4]$; there exist such zeros because y_0 lies in the support of ν which is included in $(-1, 1)$. Consider the polynomial

$$P_{n-1}(t) = \overline{q_n(t)} T(t) T_l^2(t) \Big/ \prod_{j=1}^{k+2l+1} (t - \overline{\alpha_{n,j}})$$

of degree at most $n-1$. Shrinking η if necessary, we can achieve that the argument of

$$1 \Big/ \prod_{j=1}^{k+2l+1} (t - \overline{\alpha_{n,j}})$$

is smaller than δ in absolute value for all $t \in \mathbf{R} \setminus [y_0 - \rho/2, y_0 + \rho/2]$. Since on S the argument of

$$\frac{q_n(t)}{Q_n^2(t)} \overline{q_n(t)} T(t) T_l^2(t) e^{i\varphi(t)}$$

lies in the interval $[-\pi/2 - 2\delta, \pi/2 + 2\delta]$ by the choice of T and T_l provided that $T(t) \neq 0$, it follows that the argument of

$$\frac{q_n(t)}{Q_n(t)^2} \frac{\overline{q_n(t)} T(t) T_l^2(t)}{\prod_{j=1}^{k+2l+1} (t - \overline{\alpha_{n,j}})} e^{i\varphi(t)} = \frac{q_n(t)}{Q_n(t)^2} P_{n-1}(t) e^{i\varphi(t)}$$

lies in $[-\pi/2 + \delta, \pi/2 - \delta]$ on $S \setminus [y_0 - \rho/2, y_0 + \rho/2]$, except when $T(t) = 0$. But by definition the $\alpha_{n,j}$ cannot lie on $S \setminus [y_0 - \rho/2, y_0 + \rho_2]$, hence a zero of T is necessarily a zero of P_{n-1}. Therefore (5) implies:

$$\begin{aligned} &\Re \int_{S\setminus[y_0-\rho/2,y_0+\rho/2]} \frac{q_n(t)}{Q_n(t)^2} P_{n-1}(t) e^{i\varphi(t)} d\mu(t) \\ &\geq \frac{\sin\delta}{(2\mathrm{diam}(S))^{k+2l+1}} \int_{[x_n-\rho/2,x_n+\rho/2]} \frac{|q_n(t)|^2}{|Q_n(t)|^2} |T(t)||T_l^2(t)| d\mu(t) \\ &\quad \geq \frac{\sin\delta}{(2\mathrm{diam}(S))^{k+2l+1}} e^{-2nd+n\tau}. \end{aligned} \tag{6}$$

On the other hand, the limit distribution of the zeros of the polynomials P_{n-1} for $n \in \mathbf{N}_2$ is again ν; therefore, exactly as in (4), we get

$$|P_{n-1}(y)/Q_n(y)| \leq C_1 e^{-nd}, \qquad y \in [y_0 - \rho, y_0 + \rho]$$

for sufficiently large $n \in \mathbf{N}_2$ (observe that the principle of descent that led to (4) is valid, for the zeros of P_{n-1} remain bounded, being either zeros of T, T_l which are fixed, or of $\overline{q}_n$ whose roots are in the unit disk.

We obtain from this and (4) the estimate

$$\left| \int_{y_0-\rho/2}^{y_0+r/2} \frac{q_n(t)}{Q_n^2(t)} P_{n-1}(t) d\lambda(t) \right| \leq C_2 e^{-2nd}, \tag{7}$$

and again the contradiction comes from the fact that the sum of the two integrals on left of (7) and (6) cannot be zero.

We need finally to handle the case when some zeros of Q_n go to infinity, say as $n \to \infty$, $n \in \mathcal{N}$. The number of such zeros is necessarily bounded uniformly with n, because only a bounded number of zeros of q_n can tend to 0 by Corollary (5.3) and the fact that $S \subset (0,1)$. Call these zeros $\alpha_{n,1}, \ldots, \alpha_{n,m}$, $n \in \mathcal{N}$ where m is a fixed number. Let us define

$$Q_n^*(t) = \frac{Q_n(t)}{\Pi_{j=1}^m (t - 1/\overline{\alpha}_{n,j})},$$

$$d\mu_n(t) = \left| \Pi_{j=1}^m (\overline{\alpha}_{n,j} t - 1) \right|^{-2} d\mu(t),$$

$$\varphi_n(t) = \varphi(t) - 2 \sum_{j=1}^m \operatorname{Arg}(\overline{\alpha}_{n,j} t - 1),$$

noting that all the above quantities are well defined for n large enough. Then, upon renormalizing (2), we can rewrite these orthogonality relations as:

$$\int \frac{q_n(t)}{(Q_n^*)^2(t)} t^k e^{i\varphi_n(t)} d\mu_n(t) = 0, \qquad 0 \leq k < n. \tag{8}$$

Because Q_n^* has asymptotic zero-distribution σ and its roots are now bounded, we seek to apply the first part of the proof to μ_n and φ_n instead of μ and φ even though these depend on n. As for μ_n this is possible because

$$\left| \Pi_{j=1}^m (\overline{\alpha}_{n,j} t - 1) \right| \longrightarrow 1$$

uniformly on S as $n \to \infty$, hence the constant c in (1) can be made uniform with respect to n, and the total mass remains bounded independently of n (compare the end of the proof of Theorem 3.1). As for φ_n, we notice that $\varphi_n - \varphi$ converges uniformly to zero on the convex hull of S together with all its derivatives, thus $V(\varphi - \varphi_1)$ goes to zero as well. Therefore

Corollary 5.3 remains true, and Lemma 3.4 also remains true with the same T when φ gets replaced by φ_n if n is large enough. Because these were the only facts we used on φ, this demonstrates Theorem 5.1.

■

In the previous proof we used Lemmas 5.2 and 5.4, and we need now to establish them.

Proof of Lemma 5.2. We follow the proof of Lemma 3.2. Exactly as in that lemma (remember that $1/\overline{\xi_k}$ is not present below if $\xi_k = 0$), it follows that

$$V(\varphi) + \sum_{k=1}^{n}\sum_{j=1}^{m} V\left(\arg\left(\frac{t-\xi_k}{(t-1/\overline{\xi_k})^2}\right), [a_j, b_j]\right) > (n-m+1)\pi, \tag{9}$$

otherwise there is a polynomial R_{n-1} of degree at most $n-1$ such that with some u_0

$$\Im\left\{e^{-iu_0}\frac{q_n(t)}{Q_n^2(t)}e^{i\varphi(t)}R_{n-1}(t)\right\} \tag{10}$$

has constant sign on S except at finitely many points where it vanishes. But this contradicts the fact that, by (2), we have (recall that $d\lambda(t) = e^{i\varphi(t)}d\mu(t)$):

$$\int e^{-iu_0}\frac{q_n(t)}{Q_n^2(t)}e^{i\varphi(t)}R_{n-1}(t)d\mu(t) = 0.$$

Now (9) implies what we want once we have shown that if $[a, b] \subset (0, 1)$ is any interval and if we let

$$g_\xi(t) = \arg\left(\frac{t-\xi}{(t-1/\overline{\xi})^2}\right) = \arg(t-\xi) + 2\arg(t-1/\xi)$$

where the second summand is not present if $\xi = 0$, then for $|\xi| \leq 1$ and $\xi \notin [a, b]$:

$$V(g_\xi, [a, b]) \leq \text{Angle}(\xi, [a, b]). \tag{11}$$

In fact, using the same argument as in the proof of Lemma 3.2, it is enough to prove (3) when $\xi_k \notin \cup_j[a_j, b_j]$, and then plugging (11) into (9) for each $[a, b] = [a_j, b_j]$ we get the statement in Lemma 5.2.

Note that (11) certainly holds with equality for $|\xi| = 1$, since then $g_\xi(t) = \arg(1/(t-\xi)) = -\arg(t-\xi)$ is a monotone function of $t \in [a, b]$. Another special case when (11) holds with equality is when $\xi \in [-1, 1] \setminus [a, b]$, for then both sides are 0. Thus, let $|\xi| < 1$. We may assume that ξ lies in the upper half Δ_+ of the unit disk.

For $\xi \in \Delta_+$, the function $\xi \to V(g_\xi, [a, b])$ is continuous. Moreover, by definition, $V(g_\xi, [a, b])$ is the supremum of all sums of the form

$$\Sigma(\xi) = \sum_{k=1}^{M-1} \varepsilon_k\left(g_\xi(t_{k+1}) - g_\xi(t_k)\right)$$

where $a = t_1 < t_2 < \cdots < t_M = b$ is any sequence of numbers in $[a, b]$ and $\varepsilon_k = \pm 1$ any sequence of signs. Since g_ξ is a harmonic function of ξ in Δ_+ so is every $\Sigma(\xi)$, and thus $V(g_\xi, [a, b])$, being the supremum of a family of harmonic functions and being continuous, is subharmonic. Since

$$\text{Angle}(\xi, [a, b]) = \arg(\xi - b) - \arg(\xi - a)$$

is also harmonic, it is sufficient to show that for quasi-every z on the boundary of Δ_+, we have

$$\lim_{\xi \to z} V(g_\xi, [a, b]) \le \lim_{\xi \to z} \text{Angle}(\xi, [a, b]).$$

We have seen this already on $|\xi| = 1$ and on $[-1, 1] \setminus ([a, b] \cup \{0\})$, and if ξ approaches an inner point of $[a, b]$ then both $\text{Angle}(\xi, [a, b])$ and $V(g_\xi, [a, b])$ approach π. Hence we obtain (11) via the maximum principle for subharmonic functions.

■

Proof of Lemma 5.4. The geometry is here simplified by our assumption that the support $S \subseteq [a, b]$ lies within $(0, 1)$.

It is sufficient to prove that the functions $Q_n(t)/|Q_n(t)|$ have uniformly bounded derivatives on $[a, b]$. In fact, then each of these functions can be extended to a function $\Phi_n(t)$ defined on $[0, 1]$ such that the the derivatives Φ_n' are uniformly bounded, say $|\Phi_n'| \le L$ for $t \in [0, 1]$. By Jackson's theorem [8, Theorem 6.2] there is a constant C and there are polynomials $T_{l,n}$ of degree at most l such that

$$|\Phi_n(t) - T_{n,l}(t)| \le \frac{CL}{l}, \qquad t \in [0, 1],$$

and here the right hand side is smaller than δ if $l > CL/\delta$.

In turn, it is clearly enough to verify that $\arg(Q_n(t))$ has uniformly bounded derivative on $[a, b]$. Consider a zero $1/\overline{\alpha_j} = x + iy$ of Q_n. Let the argument of this point be θ_j (which is the same as the argument of α_j), and let φ_j denote the angle in which $[a, b]$ is seen from α_j. Consider first the case where $\theta_j \in [0, \pi/4]$ and $|\alpha_j| \ge a/2$. It is elementary that θ_j is at most $\pi - \varphi_j$, and so if $\pi - \varphi_j \le \pi/4$ we get:

$$y \le (2/a) \tan \theta_j \le (2/a) \tan(\pi - \varphi_j) \le 2(2/a)(\pi - \varphi_j),$$

where we used that $\tan \gamma \le 2\gamma$ for $\gamma \in [0, \pi/4]$. If however $\pi - \varphi_j \ge \pi/4$, then

$$y \le (2/a) \le 2(2/a)(\pi - \varphi_j), \tag{12}$$

so this inequality holds regardless how large $\pi - \varphi_j$ is. In another connection, for $t \in [a, b]$,

$$\tan(\arg(1/\overline{\alpha_j} - t)) = \tan(\arg(x + iy - t)) = y/(x - t).$$

Differentiate this equality with respect to t to obtain

$$\frac{d(\arg(1/\overline{\alpha_j}-t))}{dt}\frac{1}{\cos^2(\arg(1/\overline{\alpha_j}-t))}=\frac{y}{(x-t)^2},$$

and on using (12) conclude that

$$\left|\frac{d(\arg(1/\overline{\alpha_j}-t))}{dt}\right|\leq\frac{4}{a(1-b)^2}(\pi-\varphi_j).$$

A similar estimate holds if $\theta_j\in[-\pi/4,0]$, $|\alpha_j|\geq a/2$. However, by Lemma 5.2, the number of those α_j for which either $|\alpha_j|\leq a/2$, or $\arg(\alpha_j)\notin[-\pi/4,\pi/4]$ is less than a fixed constant K_1, and since the argument of $1/\overline{\alpha_j}-t$ for each such zero clearly has bounded derivative on $[a,b]$ (with an absolute bound), the contribution to the derivative of the argument of $Q_n(t)$ of all these exceptional zeros is less than a fixed constant K_2.

So far we have proved that

$$\left|\frac{d(\arg(Q_n(t)))}{dt}\right|\leq\sum_{j=1}^{n}\left|\frac{d(\arg(1/\overline{\alpha_j}-t))}{dt}\right|\leq\frac{4}{a(1-b)^2}\sum_{j=1}^{n}(\pi-\varphi_j)+K_2.$$

Now the lemma follows from here and from Lemma 5.2.

■

6 Remarks

In this section we point out to some works that motivated the results in the preceding sections, as well as we make some remarks in connection with the method used in this paper.

1. Generally speaking, the results of the paper would hold under the hypothesis that $d\lambda/d|\lambda|$ has bounded variation while $\mu=|\lambda|$ satisfies, on its support S, the so-called Λ condition introduced in [19]:

$$\mathrm{Cap}\left\{t\in S:\ \limsup_{r\to0+}\frac{\mathrm{Log}\big(1/\mu([t-r,t+r])\big)}{\mathrm{Log}\big(1/r\big)}<+\infty\right\}=\mathrm{cap}(S)\,. \tag{1}$$

Above in all sections, we made the stronger assumption that S is regular and that μ is sufficiently thick in the sense of (1). This makes for a clearer exposition that displays already all the interesting features of the method. For the more general version, we refer the reader to [11].

2. In recent past, non-Hermitian orthogonality (see Section 3) has received a lot of interest in connection with rational approximation, for q_n in (1) is the monic denominator of the n-th diagonal Padé approximant to the Cauchy transform of λ. In this context, the main objective of this paper, namely of proving asymptotic zero distribution is tantamount to proving the convergence in capacity of these approximants [16]. As to non-Hermitian orthogonality proper, we refer the reader to [17] for a survey on the segment while [18] already deals with a more general situation where orthogonality holds over an arbitrary symmetric contour for the logarithmic potential; in this setting, [9] treats the case of a varying weight in relation to multipoint Padé approximants. Let us point out that the method in Section 3, although restricted to the segment so far, applies to measures with considerably more general support at the cost of assuming a little more on their argument, namely that it has bounded variation. It is moreover interesting in that it provides *non-asymptotic* information on the zeros of the orthogonal polynomials, *cf.* Lemma 3.2.

From the point of view of Approximation Theory, a natural sequel would be to establish convergence in capacity of best rational and meromorphic approximants on the circle to Cauchy transforms of complex measures on a segment. However, the primary purpose of this paper is to present a specific technique to handle orthogonality equations, and including such applications properties would make the paper unbalanced. These will be left here for further study.

3. The minimal degree solution of (1) is the monic denominator of the n-th diagonal Padé approximant to the Cauchy transform of λ. In this context, the possibility that $d_n < n$ accounts for cases of non-normality which are well-known to occur with such approximants, and to establish the asymptotic zero-distribution is tantamount to prove their convergence in capacity [16].

4. From the strong convergence result for Padé approximants established in [12], it follows that *all* the zeros of the minimal degree solution to (1) cluster on S when the latter is a real segment and λ is absolutely continuous with respect to Lebesgue measure with continuous nowhere vanishing (complex) density. This result is not contained in nor contains our Lemma 3.2, and it would be interesting to understand better the relations between them.

5. In connection with the Möbius transformation $\mu \to \nu$ in Section 4 (see the proof of Theorem 4.1) it is worth stressing that if M^* denotes the Markov function (1) associated to ν, then

$$M(z) = \frac{1+aw}{1-a^2} M^*(w),$$

and since $|dz| = ((1-a^2)/(1+aw)^2)|dw|$, the identity

$$\int_{\mathbf{T}} |M(z)|^2 |dz| = (1-a^2)^{-1} \int_{\mathbf{T}} |M^*(w)|^2 |dw|$$

holds. Likewise, if r_n is a rational function with numerator degree $\leq n-1$ and denominator degree n and with its poles in the unit disk and if we define a new rational function r_n^* by setting

$$r_n(z) = \frac{1+aw}{1-a^2} r_n^*(w),$$

we deduce in the same way upon regarding r_n and r_n^* as Cauchy transforms of discrete measures that

$$\int_{\mathbf{T}} |r_n(z)|^2|dz| = (1-a^2)^{-1}\int_{\mathbf{T}} |r_n^*(w)|^2|dw|.$$

Now, if we fix the denominator χ_n of r_n and if we adjust the numerator in such a way that $M - r_n$ has minimal $L^2(\mathbf{T})$-norm, we get by the characteristic property of orthogonal projections that

$$\int_{\mathbf{T}} |M(z)-r_n(z)|^2|dz| = \int_{\mathbf{T}} |M(z)|^2|dz| - \int_{\mathbf{T}} |r_n(z)|^2|dz|,$$

which is easily seen [7] to be equivalent to the vanishing of $M - r_n$ at the reflected zeros of χ_n across the unit circle. But then $M^* - r_n^*$ vanishes at the reflected zeros of the denominator of r_n^*, so that

$$\int_{\mathbf{T}} |M(z)-r_n(z)|^2|dz| = (1-a^2)^{-1}\int_{\mathbf{T}} |M^*(w)-r_n^*(w)|^2|dw|$$

and we see that best approximants also transform in a natural way.

6. As a corollary to the proof of Theorem 5.1, we get at no extra-cost a slightly extended version of this result where there is an additional weight that varies with n:

Corollary 6.1 *The conclusion of Theorem 5.1 remains true if $d\lambda$ gets replaced by $w_n d\lambda$, where w_n is a sequence of complex measurable functions on the convex hull of S whose moduli are uniformly bounded from above and below, and whose arguments are smooth with uniformly bounded derivatives.*

The corollary has special significance with respect to meromorphic approximation, as we now explain. If for $2 < p \le \infty$ we let h/q_n be a best approximant to $M(z)$ in (1) out of $H^p/\mathcal{Q}_n$, where H^p is the Hardy space of the disk and $\mathcal{Q}_n$ the polynomials of degree at most n, the (monic) denominator q_n has exact degree n, it has all its zeros in the unit disk, and it satisfies the orthogonality relations:

$$\int \frac{q_n(t)}{Q_n^2(t)} t^k\, w_n(t)\, d\lambda(t) = 0, \qquad 0 \le k < n, \tag{2}$$

where w_n is analytic without zeros in the disk; in fact, w_n is the outer factor of an n-th singular vector of the (generalized) Hankel operator with symbol $M(z)$ [3]. In that work it is shown, provided λ is analytic, that w_n is a normal family of functions which does not have the null function as an accumulation point. Hence the hypotheses of Corollary 6.1 are met and (2) implies that the asymptotic distribution of the zeros of q_n is Ω_S. Minor modifications in the arguments of [3] prove that w_n is still normal if λ is not analytic but merely satisfies the hypotheses of Theorem 2, hence Corollary 6.1 actually settles the asymptotic behaviour of the poles of the best-L^p meromorphic approximants to $M(z)$ when $2 < p \le \infty$. It is worth noting that when $p = 2$, the best meromorphic approximants to $M(z)$ are nothing but the best-L^2 rational approximants that we just considered. Moreover, a computation resembling much the one we did at the begining of the proof of Theorem 4.1 shows that all these results are valid is the support of λ is a hyperbolic geodesic arc in the disk although not necessarily a segment. The situation when $p < 2$ is still not well-understood.

Proof of Corollary 6.1. Follow the reasoning in the proof of Theorem 5.1. We can extend the definition of w_n to $[-1, 1]$ without increasing the variation nor the bound on the derivative of the argument. Upon setting $d\mu_n = |w_n| d\mu$ and $\varphi_n = \varphi + \mathrm{Arg}(w_n)$, we find ourselves in a situation like at the end of the proof of Theorem 5.1 when we had to extend the previous argument to varying φ and μ. Here again, dealing with μ_n is no problem because with our assumptions the constant c in (1) can be made uniform with respect to n, and the total mass is uniformly bounded. To handle φ_n, we notice first that it has uniformly bounded variation by the boundedness of $d\mathrm{Arg}(w_n(t))/dt$, hence Corollary 5.3 continues to hold. Second, using Jackson's theorem as in the proof of Lemma 5.4 instead of the Weierstrass theorem, we find that Lemma 3.4 will hold with φ_n instead of φ and T_n instead of T, where the degree of T_n is uniformly bounded. From the way T was constructed in that lemma, $|T_n|$ will also be uniformly bounded (note that the factors needed to handle the discontinuities of φ will not depend on n). We can thus assume, up to extracting a subsequence, that T_n converges and thus can be made independent of n when the latter is large enough (see the end of the proof of Theorem 5.1). This is all we need to carry out the proof as in Theorem 5.1.

■

References

[1] J. E. Andersson, Best rational approximation to Markov functions, *J. Approx. Theory*, 76, 219–232, 1994.

[2] L. Baratchart and V. Prokhorov and E.B. Saff, Best meromorphic approximation of Markov functions on the unit circle, *Foundations of Constructive Math*, 1 (4), 385-416, 2001.

[3] L. Baratchart and F. Seyfert, An L^p analog to AAK theory for $p \geq 2$, *Journal of Functional Analysis*, 191, 52-122, 2002.

[4] L. Baratchart and H. Stahl and F. Wielonsky, Non-uniqueness of rational best approximants, *J. of Comp. and Applied Math.*, Proc. of Conference on Continuous Fractions, in honour of H. Waadeland, Trondheim, 1997.

[5] L. Baratchart and H. Stahl and F. Wielonsky, Asymptotic error estimates for L^2 best rational approximants to Markov functions on the unit circle, *Journal of Approximation Theory*, 108, 53-96, 2001.

[6] L. Baratchart and H. Stahl and F. Wielonsky, Asymptotic uniqueness of best rational approximants of given degree to Markov functions in L^2 of the circle, *Constructive Approximation*, 17, 103–138, 2001.

[7] L. Baratchart and F. Wielonsky, Rational approximation in the real Hardy space H^2 and Stieltjes integrals: a uniqueness theorem, *Constructive Approximation*, 9, pp. 1-21, 1993.

[8] R. A. DeVore and G. G. Lorentz, Constructive Approximation, *Grundlehre der mathematischen Wissenschaften*, **303**, Springer Verlag, New York 1993.

[9] A.A. Gonchar and E.A. Rakhmanov, Equilibrium distributions and the degree of rational approximation of analytic functions, *Math. USSR Sbornik*, 176, pp. 306–352, 1989.

[10] R. Kannan and C.K. Krueger, *Advanced Analysis on the real line*, Universitext, Springer-Verlag, New-York, 1996.

[11] R. Küstner, *Distribution Asymptotique des zéros de polynômes orthogonaux par rapport à des mesures complexes ayant un argument à variation bornée*, Ph.D. thesis, University of Nice, 2003.

[12] A. Magnus, *Toeplitz matrix techniques and convergence of complex weight Padé approximants*, J. Computational and Applied Math, 19, 23–38, 1987.

[13] Q.I. Rahman and G. Schmeisser, *Analytic Theory of Polynomials*, London Math. Soc. monographs, 2002.

[14] T. Ransford, *Potential Theory in the Complex Plane*, Cambridge University Press, Cambridge, 1995

[15] E. B. Saff and V. Totik, *Logarithmic Potentials with External Fields*, Grundlehren der mathematischen Wissenschaften, **316**, Springer Verlag, Berlin, Heidelberg, 1997.

[16] H. Stahl, The convergence of Padé approximants to functions with branch points, *J. of Approximation Theory*, 91, 139–204, 1997.

[17] H. Stahl, Orthogonal polynomials with respect to complex-valued measures, *Orthogonal Polynomials and their applications*, C. Brezinski and L. Gori and A. Ronveaux *eds*, J.C. Baltzer AG, Scientific Publishing Co., IMACS, pp. 139-154, 1991.

[18] H. Stahl, Orthogonal polynomial with complex-valued weight functions I,II. *Constructive Approx.*, 2, pp. 225–240 and 241–251, 1996.

[19] H. Stahl and V. Totik *General Orthogonal Polynomials* Encyclopedia of Mathematics, **43**, Cambridge University Press, New York 1992

[20] J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, Colloquium Publications, **20**, Amer. Math. Soc., Providence, 1960

Reinhold Küstner
U.F.R. de Mathématiques
Université des Sciences et Technologies de Lille 1, Cité Scientifique,
59655 Villeneuve d'Ascq Cedex
France
Reinhold.Kustner@math.univ-lille1.fr

Laurent Baratchart
INRIA
2004 route des Lucioles, BP 93, 06902
Sophia-Antipolis Cedex
France
Laurent.Baratchart@sophia.inria.fr

Vilmos Totik
Bolyai Institute
University of Szeged
Aradi v. tere 1,
6720, Hungary and

Department of Mathematics
University of South Florida
Tampa, FL 33620
USA
totik@math.usf.edu

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399